True friends or false friends?
Lexical similarity for predicting cross-signing success

Carl (Calle) Börstell
Onno Crasborn
Lori Whynot

Radboud University
Nijmegen, the Netherlands

Communication across languages

- Two people meet.
  They don't know each other.
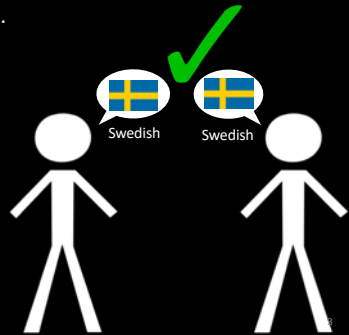 They don't have a shared language.

How do they communicate?

Communication across languages

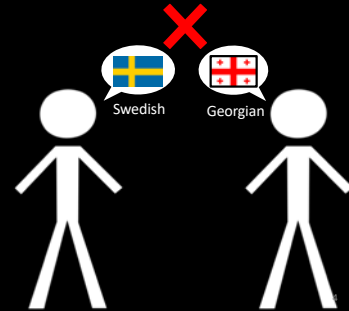- If they *do* speak the same language…

Swedish   Swedish

No problem!

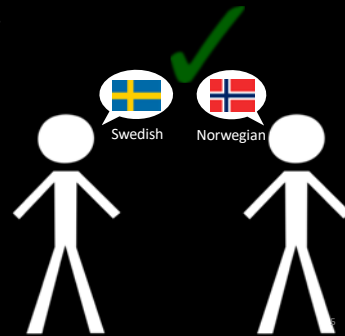## Communication across languages

- If they speak the different languages…

Swedish   Georgian

Problematic!

Communication across languages

- If they speak the _similar_ languages…

Swedish   Norwegian

It depends!

True friends & false friends

- Some word forms are similar across languages…

- If they mean the _same_ thing…

  = **TRUE FRIENDS**!

  teckenspråk tegnspråk
  'sign language' 'sign language'

- If they mean _different_ things…

  = **FALSE FRIENDS**

  rolig rolig
  'funny' 'calm'

6

We can communicate across some languages, like Swedish and Norwegian
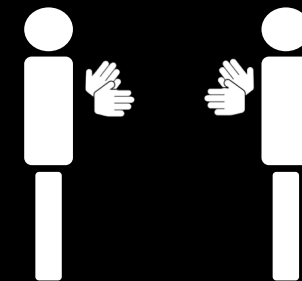
… but what about sign languages?

7

Cross-signing

- Previous research points to communication across different languages being possible – _cross-signing_
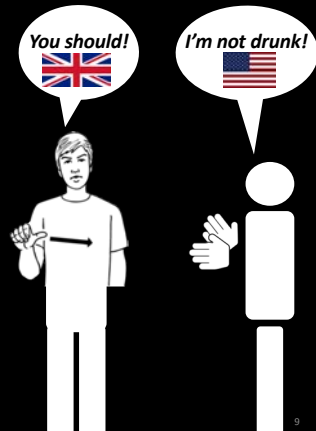
  … even when the two sign languages are unrelated!

Supalla & Webb 1995; Zeshan 2015; Byun et al. 2018; _inter alia_

8

Slide 9

True or false friends?
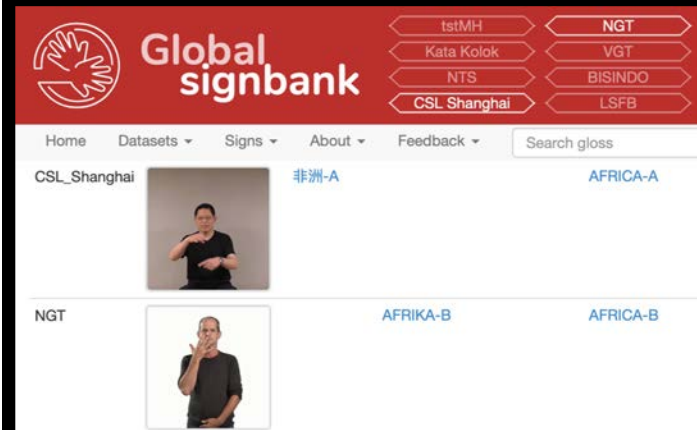
You should!

I'm not drunk!

- If you see a sign from another SL:

  - You know the sign
    - L2 knowledge ✓

  - The meaning is the same in your SL ✓
    - True friend

  - The meaning is different in your SL ✗
    - False friend

McKee et al. 2019

9

Slide 10

Can we use true/false friends to measure linguistic distance?

... If so, does it predict cross-signing success between those languages?

10

Slide 11

Global signbank

tstMH	NGT
Kata Kolok	VGT
NTS	BISINDO
CSL Shanghai	LSFB

Home   Datasets ▾   Signs ▾   About ▾   Feedback ▾   Search gloss

CSL_Shanghai   非洲-A   AFRICA-A

NGT   AFRIKA-B   AFRICA-B

Phonological properties

⬇

Two-handed, neutral space, ...

⬆ **Compare these!** ⬇

One-handed, mouth, ...

11

Slide 12

Example: Comparing signs for 'no'

NGT   **CSL**

12

Example: Comparing signs for 'no'

NGT   CSL

	No of hands	Handshape	Location	Mov. direction	Repeated mov.	TOTAL
NGT	1					
CSL	1					
Similarity	1					

Example: Comparing signs for 'no'

NGT   CSL

	No of hands	Handshape	Location	Mov. direction	Repeated mov.	TOTAL
NGT	1	B				
CSL	1	B				
Similarity	1	1				

Example: Comparing signs for 'no'

NGT   CSL

	No of hands	Handshape	Location	Mov. direction	Repeated mov.	TOTAL
NGT	1	B	Neutral			
CSL	1	B	Neutral			
Similarity	1	1	1			

Example: Comparing signs for 'no'

NGT   CSL

	No of hands	Handshape	Location	Mov. direction	Repeated mov.	TOTAL
NGT	1	B	Neutral	Ipsilateral		
CSL	1	B	Neutral	Ipsilateral		
Similarity	1	1	1	1		

Example: Comparing signs for 'no'

NGT   **CSL**

	No of hands	Handshape	Location	Mov. direction	Repeated mov.	TOTAL
NGT	1	B	Neutral	Ipsilateral	No	
CSL	1	B	Neutral	Ipsilateral	Yes	
Similarity	1	1	1	1	0	

Example: Comparing signs for 'no'

NGT   **CSL**

	No of hands	Handshape	Location	Mov. direction	Repeated mov.	TOTAL
NGT	1	B	Neutral	Ipsilateral	No	
CSL	1	B	Neutral	Ipsilateral	Yes	
Similarity	1	1	1	1	0	4/5 = 0.8

Phonological properties in Signbank

- *Handedness*
- *Strong Hand*
- *Weak Hand*
- *Handshape Change*
- *Relation between Articulators*
- *Location*
- *Relative Orientation: Movement*

- *Relative Orientation: Location*
- *Orientation Change*
- *Contact Type*
- *Movement Shape*
- *Movement Direction*
- *Repeated Movement*
- *Alternating Movement*

19

Languages in Global Signbank

- Sign Language of the Netherlands (NGT)
  - 3,531 coded signs

- Shanghai Chinese Sign Language (CSL)
  - 568 coded signs

- International Sign (IS)
  - 200 coded signs
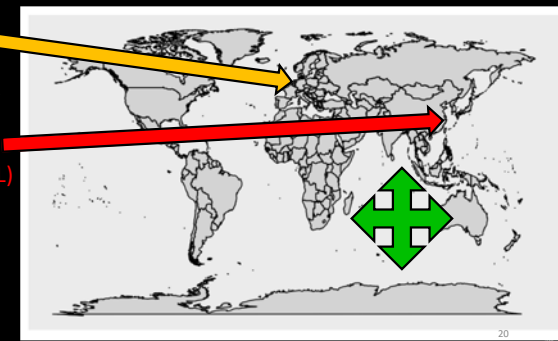
20

Finding our friends

A short Python script →

1. Compare all FORMS
   – every sign in both datasets

2. Compare matched MEANINGS
   – only meaning-matched signs

Language A — Similarity score — **Language B**

Language A	Similarity score	Language B
SIGN-A	0.88	SIGN-A
SIGN-B	0.33	SIGN-B
SIGN-C	0.56	SIGN-C
SIGN-D	0.36 ✓ 1.00	SIGN-D

Comparing signs, part 1

From FORM to MEANING

23

Finding our friends

GOOD
GOOD

- Comparing NGT and Shanghai CSL

- Script gives us 30 *sign pairs* that have identical forms in both languages

BEGIN

- A manual check proves that
  - 12 pairs are true friends
  - 18 are false friends

TO-LIFT

Finding our friends

- Comparing NGT and IS

- Script gives us 10 *sign pairs* that have identical forms in both languages

- A manual check proves that
  - 6 pairs are true friends
  - 4 are false friends

ZERO
ZERO
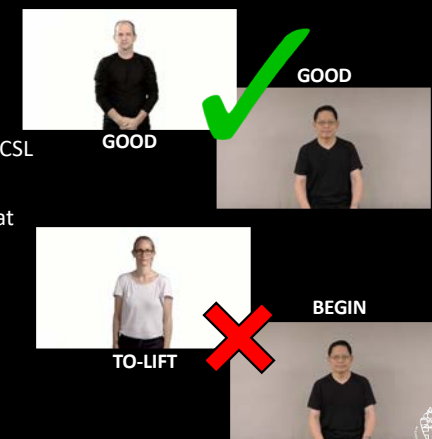ONE
TUESDAY

Finding our friends

- Comparing CSL and IS

- Script gives us 0 *sign pairs* that have identical forms in both languages

- But the two datasets are small (568 signs ~ 200 signs)

?

Interim findings: true/false friends

- Our method helps us identify true/false friends across languages of Global Signbank semi-automatically

- The number of friends found depends on size of datasets (languages)

- Are NGT~IS closer than NGT~CSL? Too little data!

Comparing signs, part 2
From MEANING to FORM

Lexical similarity

- Linking specific CONCEPTS, we can measure how similar two languages are

- A script automatically matches *sign glosses* to the 3,431 *concepts* in the Concepticon database:

  1. CONTEMPTIBLE
  2. DUST
  3. BRAVE
  4. COURTYARD
  5. GAZELLE
     …

Concepticon — Home | Concepts | Concept sets | Concept lists | Compilers

Welcome to the Concepticon

This resource presents an attempt to link the large amount of different concept lists which are used in the linguistic literature, ranging from Swadesh lists in historical linguistics to naming tests in clinical studies and psycholinguistics.
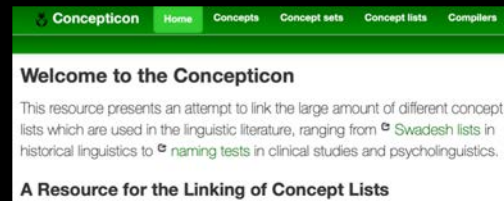
A Resource for the Linking of Concept Lists

List et al. 2019

29

Lexical similarity script

Signbank — NGT ← Comparison → CSL

Concepticon — 'no'
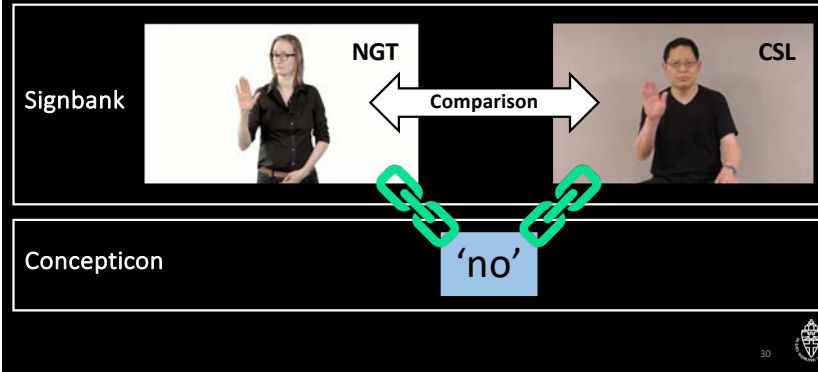
30

Lexicostatistics

Previous methods
- Only basic form parameters:
  - Handshape
  - Location
  - Movement
  - Orientation

- 3/4 or 2/3 counts as "similar"

Our methods
- Looks at more detailed phonological properties

- Automatically with a computer; uniform database

- Gives a more precise score (0–1)

Woodward 1991, 1993, 2000; Bickford 2005; Al-Fityani & Padden 2010; *inter alia*

31

Lexical similarity

- Very few signs are *true* friends or *near* friends

32

Lexical similarity: matches

- Looking at the top form-similar matched concepts, we find
  - *Numbers*: ZERO, ONE, TWO, …
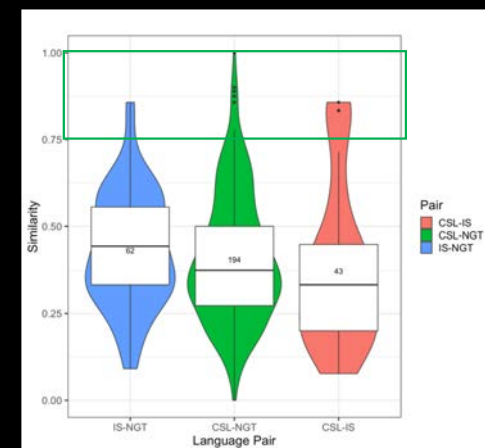 - *Body-parts*: FACE, ARM, EAR
  - *Gestures*(?): GOOD, NO, MONEY

- We need more data!

CONCEPT	Pair	Score
GOOD	CSL-NGT	1.00
HAPPY	CSL-NGT	0.90
ELECTRICITY	CSL-NGT	0.89
NO	CSL-NGT	0.88
GOOD	IS-NGT	0.86
FACE	CSL-NGT	0.86
GOOD	CSL-IS	0.86
MONEY	IS-NGT	0.83
ONE	CSL-IS	0.83
FOUR	CSL-IS	0.83
ZERO	CSL-IS	0.83
TWO	CSL-IS	0.83
COLD	CSL-NGT	0.78
ARM	CSL-NGT	0.78
EAR	CSL-NGT	0.75
WHAT	CSL-NGT	0.75

Conclusions

- A two-part method for comparing lexical similarity across languages of Global Signbank

- Both methods suggest that CSL is more distinct from NGT and IS
  - Supports recent cross-linguistic work on phonology of Western vs. Eastern SLs – but our datasets are still small

- We will use these metrics in our research on communicative success in cross-signing

Yu et al. 2018

34

THANK YOU!

This study was funded by the *Netherlands Organisation for Scientific Research (NWO)* grant number 277-70-014.

We would like to acknowledge the rest of our team, Tashi Bradford, Aurélia Nana Gassa Gonga, Maya de Wit, and Merel van Zuilen. A special thanks to Anique Schüller and Neil Ray for annotating specific Signbank data for this study.

Radboud University
Nijmegen, the Netherlands

35

References

- Al-Fityani, Kinda & Carol Padden. 2010. Sign languages in the Arab world. In Diane Brentari (ed.), Sign languages: A Cambridge language survey, 433–450. New York, NY: Cambridge University Press.
- Bickford, J. Albert. 2005. The signed languages of Eastern Europe. SIL International & University of North Dakota. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.103.5039&rep=rep1&type=pdf.
- Byun, Kang-Suk, Connie de Vos, Anastasia Bradford, Ulrike Zeshan & Stephen C. Levinson. 2018. First Encounters: Repair Sequences in Cross-Signing. Topics in Cognitive Science 10(2). 314–334. doi:10.1111/tops.12303.
- Hiddinga, Anja & Onno Crasborn. 2011. Signed languages and globalization. Language in Society 40(4). 483–505. doi:10.1017/S0047404511000480.
- McKee, David, Rachel Locker McKee, Sara Pivac Alexander, Lynette Pivac & Mireille Vale. 2019. Online Dictionary of New Zealand Sign Language. Deaf Studies Research Unit, Victoria University of Wellington. http://nzsl.vuw.ac.nz/.
- R Core Team. 2017. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.
- Sáfár, Anna, Laurence Meurant, Thierry Haesenne, Ellen Nauta, Danny De Weerdt & Ellen Ormel. 2015. Mutual intelligibility among the sign languages of Belgium and the Netherlands. Linguistics 53(2). 353–374. doi:10.1515/ling-2015-0004.
- Supalla, Ted & Rebecca Webb. 1995. The grammar of international sign: A new look at pidgin languages. In Karen Emmorey & Judy S. Reilly (eds.), Language, Gesture, and Space, 333–352. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Wickham, Hadley. 2016. ggplot2: Elegant Graphics for Data Analysis. New York, NY: Springer-Verlag. http://ggplot2.org.
- Woodward, James. 1991. Sign language varieties in Costa Rica. Sign Language Studies 73. 329–346.
- Woodward, James. 1993. The relationship of sign language varieties in India, Pakistan, and Nepal. Sign Language Studies 78. 15–22.
- Woodward, Jr., James C. 2000. Sign language and sign language families in Thailand and Viet Nam. In Karen Emmorey & Harlan Lane (eds.), The signs of language revisited: An anthology to honor Ursula Bellugi and Edward Klima, 23–47. Mahwah, NJ: Lawrence Erlbaum Associates.
- Yu, Shi, Carlo Geraci & Natasha Abner. 2018. Sign Languages and the Online World Online Dictionaries & Lexicostatistics. In Nicoletta Calzolari, Khalid Choukri, Christopher Cieri, Thierry Declerck, Sara Goggi, Koiti Hasida, Hitoshi Isahara, et al. (eds.), Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018). Miyazaki, Japan: European Language Resources Association (ELRA).
- Zeshan, Ulrike. 2015. "Making meaning": Communication between sign language users without a shared language. Cognitive Linguistics 26(2). 211–260. doi:10.1515/cog-2015-0011.

36

9